How to use Hue with Kylin

Date: August 2016

Author: Ramón Portolés, Alberto a.ramonportoles@gmail.com Linkedin

Intro

In Hue-2745 v3.10, add jdbc support like Phoenix, **Kylin**, Redshift, Solr Parallel SQL, …

However, there isn't any manual to use with Kylin :(

Pre-requisites

Build a cube sample of Kylin with Quick Start with Sample Cube,

You can check:

Used Software:

- Apache Hue v3.10.0
- Apache Kylin v1.5.2

Install Apache Hue

If you have Hue installed, you can skip this step

I installed Hue on my Ubuntu 16.04 LTS. The official Instructions didn't work but this works fine:

There isn't any binary package, thus we need to install all pre-requisites to compile with the command *make*

```
Sudo apt-get install --fix-missing -q -y \
  git \
  ant \
  gcc \
  g++ \
  libkrb5-dev \
  libmysqlclient-dev \
```

```
libssl-dev \
libsasl2-dev \
libsasl2-modules-gssapi-mit \
libsqlite3-dev \
libtidy-0.99-0 \
libxml2-dev \
libxslt-dev \
libffi-dev \
make \
maven \
libldap2-dev \
python-dev \
python-setuptools \
libgmp3-dev \
libz-dev
```

Download and Compile:

> *git clone https://github.com/cloudera/hue.git*
> *cd hue*
> *make apps*

Start and connect to Hue:

> *build/env/bin/hue runserver_plus localhost:8888*
> - runserver_plus: is like runserver with debugger
> - localIP: Port, usually Hue uses 8888

The output must be similar to:


```
Django version 1.6.10, using settings 'desktop.settings'
Development server is running at http://localhost:8888/
Using the Werkzeug debugger (http://werkzeug.pocoo.org/)
```

Connect using your browser: http://localhost:8888

Since this is your first time logging in, pick any
username and password. Be sure to remember
these, as **they will become your Hue superuser
credentials.**

Username

Password

Create Account

Important: The first time that you connect to hue, you set Login / Pass for admin

We will use Hue / Hue as login / pass

Issue 1: Could not create home directory

Could not create home directory.

It is a permission problem of your current user, you can use: *sudo* to start Hue

Issue 2: Could not connect to …

Could not connect to localhost:10000

If you have downloaded Hue's code from Git, Hive connection is active but not configured → you
can skip this message

Issue 3: Address already in use

[Errno 98] Address already in use

The port is in use or you have a Hive process running already.

You can use *ps -ef | grep hue,* to find the PID and kill

Configure Apache Hue for Kylin

Our purpose is to add a snipped in a notebook with Kylin queries

References:

- [Custom SQL Databases](#)

- [Manual: Kylin JDBC Driver](#)

- [GitHub: Kylin JDBC Driver](#)

Register JDBC Driver

1. We need to find the JAR Class for the JDBC Connector
   From [Kylin Download](#), Choose **Binary** and the **correct version of Kylin and HBase**
   Download & Unpack: in ./lib:

2. Place this JAR in Java ClassPATH using .bashrc

   `export CLASSPATH=~/Descargas/apache-kylin-1.5.2-bin/lib/kylin-jdbc-1.5.2.jar`

   check the actual value: `echo $CLASSPATH`

   check the permission for this file (must be accessible to you):
   `-rwxr-xr-x 1 root root 11640840 ago 24 23:26 kylin-jdbc-1.5.2.jar`

3. Add this new interface to Hue.ini
   Where is the hue.ini ?
   - If the code is downloaded from Git : *UnzipPath/desktop/conf/pseudo-distributed.ini*
     (I shared my *INI* file in GitHub)
   - If you are using Cloudera: you must use Advanced Configuration Snippet
   - Other: find your actual *hue.ini*

   Add these lines in [[interpreters]]

```
 [[[kylin]]]
name=kylin JDBC
interface=jdbc
options='{"url": "jdbc:kylin://172.17.0.2:7070/learn_kylin","driver": "org.apache.kylin.jdbc.Driver", "user":
        "ADMIN", "password": "KYLIN"}'
```
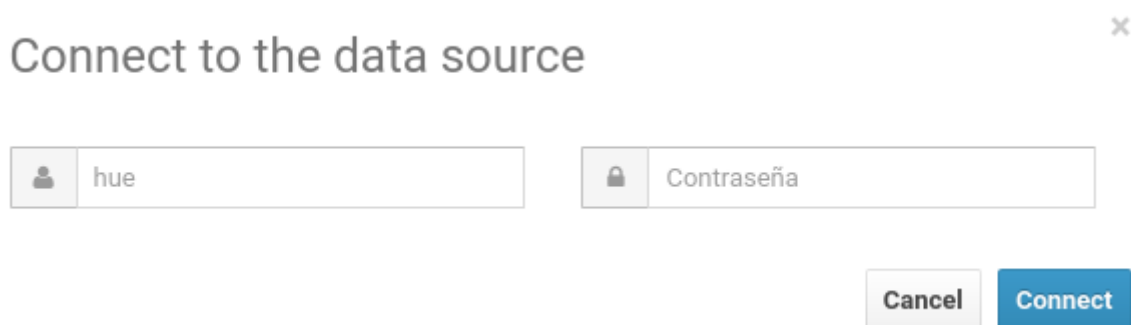
4. Try to Start Hue and connect, just like in 'Error: Reference source not found'

TIP: you will need one JDBC Source for each project

Register without a password, It can do use this other format:

```
options='{"url": "jdbc:kylin://172.17.0.2:7070/learn_kylin","driver": "org.apache.kylin.jdbc.Driver"}'
```

And when you open the Notebook, Hue prompts this:

Issue 1: Hue can't Start

If you see this when you connect to Hue (http://localhost:8888):

You can go to the last line: `obj, end = self.scan_once(s, idx)`

And launch Python Interpreter (see console icon on the right):

```
[console ready]
>>> print (s)
{"url": "jdbc:kylin://172.17.0.2:8080/learn_kylin, "driver": "org.apache.kylin.jdbc.Driver", "user":
"ADMIN", "password": "KYLIN"}
>>>
```

In my case: I've forgotten to close " after learn_kylin

Issue 2: Password Prompting

In Hue 3.11 there is a bug: Hue 4716

In Hue 3.10 with Kylin, I didn't have any problem   :)

Test query Example

Add Kylin JDBC as source in the Kylin's notebook:

Notebooks > + > ... > 🗄 kylin JDBC

Write a query, like this:
select part_dt, sum(price) as total_selled, count(distinct seller_id) as sellers from kylin_sales group by part_dt order by part_dt

And Execute with: ▶

```
select part_dt, sum(price) as total_selled, count(distinct seller_id) as sellers
```

	PART_DT	TOTAL_SELLED	SELLERS
1	2012-01-01	466.9037	12
2	2012-01-02	970.2347	17
3	2012-01-03	917.4138	14
4	2012-01-04	553.0541	10
5	2012-01-05	732.9007	18

Congratulations !!! you are connected to Hue with Kylin

Issue 1: No suitable driver found for jdbc:kylin

```
An error occurred while calling z:java.sql.DriverManager.getConnection. :
java.sql.SQLException: No suitable driver found for jdbc:kylin://172.17.0.2:7070/learn_kylin
at java.sql.DriverManager.getConnection(DriverManager.java:689) at
java.sql.DriverManager.getConnection(DriverManager.java:247) at
sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method) at
sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62) at
```

There is a bug (not solved since 27 Aug. 2016, nor in 3.10 and 3.11), but the solution is very easy (Thanks to Shahab Tajik)

You only need to change 3 lines in *<HuePath>/desktop/libs/librdbms/src/librdbms/jdbc.py*

Limits

Nowadays, in Hue 3.10 and 3.11

- Auto-complete doesn't work on JDBC interfaces

- Max 1.000 records. There is a limitation on JDBC interfaces, because Hue does not support result pagination Hue 3419

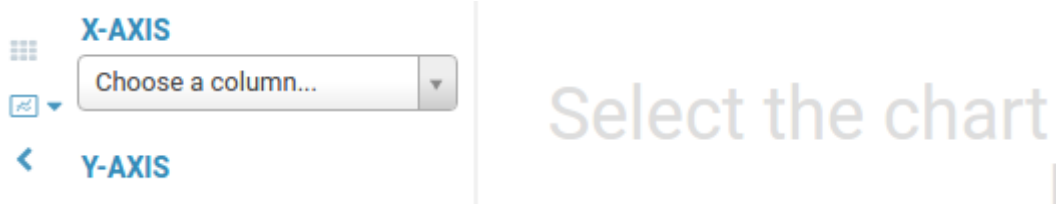- Obviously: It's read-only

Future Work

Dashboards

There is an amazing feature of Hue: Searh Dasboards / Dynamic Dashboards (you can 'play' with this Demo On-line) … **But this only works with SolR**

There is a Jira to solve this: Hue 3228 is in RoadMap for 4.1. Check Hue MailList and add Dashboards to JDBC connections

Chart & Dynamic Filter

Nowadays, it isn't compatible, you only can work with Grid

I think the problem is that Hue doesn't know how to read Kylin Metadata Hue 4011

DB Query

 > , is not yet supported with JDBC

Now it only supports Django Databases

For any suggestions, feel free to contact me

Thanks,

Alberto.